ppl.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/675,637	09/29/2000	Kenji Yamanishi	13931	1719

23389 7590 05/20/2004

SCULLY SCOTT MURPHY & PRESSER, PC
400 GARDEN CITY PLAZA
GARDEN CITY, NY 11530

EXAMINER	
SHARON, AYAL I	

ART UNIT	PAPER NUMBER
2123	7

DATE MAILED: 05/20/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PPG

Office Action Summary	Application No.	Applicant(s)	
	09/675,637	YAMANISHI ET AL.	
	Examiner	Art Unit	
	Ayal I Sharon	2123	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 29 September 2000.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-16 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-16 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☒ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date 3,6.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

DETAILED ACTION

Introduction

1. Claims 1-16 of U.S. Application 09/675,637 filed on 09/29/2000 are presented for examination.

Oath/Declaration

2. Examiner has located an article that was co-authored by the applicants: "On-line Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms." Proc. 6th ACM SIGKDD. Aug, 2000. pp.320-324. The publication post-dates the priority date of the application by 11 months. However, the article appears to read upon the claimed invention, and has a third co-author, Mr. Graham Williams, and also a fourth co-author, Mr. Peter Milne. Examiner requests clarification regarding Mr. Williams's and Mr. Milne's relationship to the current application.

Claim Objections

3. Claims 1, 6, 10, and 14 are objected to because of the following informalities: the awkward phrasing of the claims makes it difficult to determine whether the claims are method claims or device claims. The information is hidden in the

preambles. Claims 2-5, 7-9, 11-13, and 15-16, on the other hand, are understandable. Appropriate correction is required.

Double Patenting

4. Claim 1 is provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claim 1 of copending Application No. 10/179,374. Although the conflicting claims are not identical, they are not patentably distinct from each other because Claim 1 of the instant application is narrower than Claim 1 of the co-pending application. The broader claim therefore reads upon the narrower claim.
5. Claim 1 is provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claim 2 of copending Application No. 10/619,626. Although the conflicting claims are not identical, they are not patentably distinct from each other because Claim 1 of the instant application is narrower than Claim 2 of the co-pending application. The broader claim therefore reads upon the narrower claim.
6. These are provisional obviousness-type double patenting rejection because the conflicting claims have not in fact been patented.

Claim Rejections - 35 USC § 101

7. 35 U.S.C. 101 reads as follows:

   Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

8. An invention which is eligible for patenting under 35 U.S.C. § 101 is in the "useful arts" when it is a machine, manufacture, process or composition of matter, which produces a concrete, tangible, and useful result. *The fundamental test for patent eligibility is thus to determine whether the claimed invention produces a* ***"useful, concrete and tangible result."*** The test for practical application as applied by the examiner involves the determination of the following factors:

   a. "Useful" - The Supreme Court in *Diamond v. Diehr* requires that the examiner look at the claimed invention as a whole and compare any asserted utility with the claimed invention to determine whether the asserted utility is accomplished. Applying utility case law the examiner will note that:

      - the utility need not be expressly recited in the claims, rather it may be inferred.
      - if the utility is not asserted in the written description, then it must be well established.

   b. "Tangible" - Applying *In re Warmerdam*, 33 F.3d 1354, 31 USPQ2d 1754 (Fed. Cir. 1994), the examiner will determine whether there is simply a mathematical construct claimed, such as a disembodied data structure

and method of making it. If so, the claim involves no more than a manipulation of an abstract idea and therefore, is nonstatutory under 35 U.S.C. § 101. In *Warmerdam* the abstract idea of a data structure became capable of producing a useful result when it was fixed in a tangible medium which enabled its functionality to be realized.

c. "Concrete" - Another consideration is whether the invention produces a "concrete" result. Usually, this question arises when a result cannot be assured. An appropriate rejection under 35 U.S.C. § 101 should be accompanied by a lack of enablement rejection, because the invention cannot operate as intended without undue experimentation.

9. The Examiner respectfully submits, under current PTO practice, and in view of the 112(1) rejections, that the claimed invention does not recite *either a useful or tangible result.*
   a. Claims 1-3 and 5-16 calculate "a degree of outlier" for a set of data. However, it is not clear what is the practical utility of this result. Moreover, this appears to be merely a mathematical construct.
   b. Claims 4 estimate "a probability distribution of generation" of data. However, it is not clear what is the practical utility of this result. Moreover, this appears to be merely a mathematical construct.
   c. Claims 1-16 do not specifically claim a hardware implementation nor software implemented in a device or a computer-readable medium. The claims are therefore directed to a mathematical construct. They are therefore intangible according to *In re Warmerdam*.

10. **Claims 1-16 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.** The Examiner respectfully submits that the claims are directed towards intangible subject matter.

11. **Claims 1-16 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility.** The Examiner respectfully submits that Applicant's have not specifically claimed a practical application.

Claim Rejections - 35 USC § 112

12. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

13. Claims 1-16 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

Claim Rejections - 35 USC § 102

14. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

15. The prior art used for these rejections is as follows:

16. Burge, P. and Shawne-Taylor, J. "Detecting Cellular Fraud Using Adaptive Prototypes". Proc. of AI Approaches to Fraud Detection and Risk Management. Pp.72-77, 1997. (Henceforth "**Burge**").

17. Yamanishi, K. et al. "On-line Unsupervised Outlier Detection Using Finite Mixtures With Discounting Learning Algorithms." Proc. of the 6th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Pp.320-324. 2000. (Henceforth "**Yamanishi et al.**").

18. The Yamanishi et al. reference, which post-dates the foreign priority date of the application, is relevant in regards to its discussion of the Burge reference (See MPEP §2128 and *In re Epstein*, 32 F.3d 1559, 31 USPQ2d 1817 (Fed. Cir. 1994)). The Yamanishi et al. reference (See p.320, col.2, para. 3) teaches the following about the Burge reference:

> Note that there exists only a few works (e.g. Burge) focusing on the on-line unsupervised learning based approach [to outlier detection in data mining].

and also the following (See p.321, col.1, para.5) about the Burge reference:

> The design of SS [SmartSifter] was inspired by the work by Burge and Shawe-Taylor. Our work differs from [Burge] in the following regards:
>
> 1) SS [SmartSifter] treats both categorical and continuous variables, while [Burge] deals only with continuous ones.
>
> 2) While Burge uses two models in the algorithm: the long term model and the short term one, SS [SmartSifter] unifies them into one model with the aim of a clearer statistical meaning and a lower computational cost.
>
> 3) SS [SmartSifter] uses either a parametric representation for a probabilistic model or a non-parametric one, while only a non-parametric

> one is used in [Burge]. In Sec.3.1, we compare our parametric method with the non-parametric one to show that the former outperforms the latter both in accuracy and computation costs.

Examiner notes that two of the three co-authors of the article are the inventors in the present application.

19. Applicant's own admission (Specification, p.3, paragraphs 2-3) says the following about the Burge reference:

> The method by P. Burge and J. Shawe-Taylor relates to a similar fraud detection based on unsupervised data. This method, however, conducts fraud detection with two non-parametric models, a short-term model and a long-term model, to make a distance between them as a criterion for an outlier. Statistical basis of the short-term model and the long-term model is insufficient to make statistical significance of a distance therebetween [sic] unclear.
>
> In addition, preparation of two models, short-term and long-terms [sic], deteriorates calculation efficiency. Further problems are involved such as a problem that only continuous value data can be handled and not categorical data and a problem that since only non-parametric models are handled, fraud detection is unstable and inefficient.

20. Examiner notes that:

    a. The model in the Yamanishi et al. reference maps to the model in the current application - for example, compare the following equations:

        - Equation in Specification, p.24, line 5, to Equation in Burge, p.321, col.2, "Gaussian Mixture Model", net-to-last equation.
        - Equation in Specification, p.24, line 8, to Equation in Burge, p.321, col.2, last equation.
        - Equations in Specification, p.26, to Equations in Burge, p.322, col.2, "SDEM Algorithm".

- Equation in Specification, p.29, line 15 to Equation in Burge, p.322, col.2, "kernel mixture model" Eq.3.
- Equation In Specification, p.39, line 16, to Equation in Burge, p.323, col.1, "logarithmic loss", last equation.

b. The Applicants have admitted (Specification, p.3, paragraph 2) that the Burge reference "... relates a similar fraud detection based on unsupervised data ...", and the Yamanishi et al. reference teaches that the model that it discloses " ... was inspired by the work by Burge and Shawe-Taylor", and

c. The Yamanishi et al. specifies three differences between the model taught in the Burge reference and the model taught in the Yamanishi et al. reference, and

d. None of the claims in the current application recite <u>any</u> of the three stated differences from the Burge reference (that are taught by Yamanishi et al.).

e. Examiner therefore finds the current claims to be anticipated by the Burge reference.

21. The claim rejections are hereby summarized for Applicant's convenience. The detailed rejections follow.

22. Claims 1-16 are rejected under 35 U.S.C. 102(b) as being clearly anticipated by Burge.

23. In regards to Claim 1, Burge teaches the following limitations:

1. For use in a degree of outlier calculation device
for sequentially calculating a degree of outlier of each
data with a data sequence of real vector values as input,

a probability density estimation device for,
while sequentially reading said data sequence, estimating a probability distribution of generation of the data in question by using a finite mixture distribution of normal distributions, comprising:
(Burge, especially: "Protoyping")

probability calculation means for calculating, based on a value of input data and values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities, a probability of generation of the input data in question from each normal distribution; and
(Burge, especially: "Constructing Profiles")

parameter rewriting means for updating and rewriting the stored parameter values while forgetting past data, according to newly read data based on a probability obtained by the probability calculation means, values of a mean parameter and a variance parameter of each normal distribution and a weighting parameter of each normal distribution.
(Burge, especially: "The Fraud Engine")

24. In regards to Claim 2, Burge teaches the following limitations:

2. The probability density estimation device as set forth in claim 1, further comprising

parameter storage means for storing values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities and a weighting parameter of each normal distribution, wherein
(Burge, especially: "Constructing Profiles")

said parameter rewriting means updates and rewrites data of said parameter storage means.
(Burge, especially: "Constructing Profiles")

25. In regards to Claim 3, Burge teaches the following limitations:

3. A degree of outlier calculation device for sequentially calculating a degree of outlier of each data with a data sequence of real vector values as input, comprising:

a probability density estimation device for, while sequentially reading said data sequence, estimating a probability distribution of generation of the data in question by using a finite mixture of normal distributions including
(Burge, especially: "Constructing Profiles")

(a) parameter storage means for storing values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities and a weighting parameter of each normal distribution,

(Burge, especially: "Constructing Profiles")

(b) probability calculation means for calculating, based on a value of input data and values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities, a probability of generation of the input data in question from each normal distribution, and

(Burge, especially: "Constructing Profiles")

(c) parameter rewriting means for updating and rewriting the stored parameter values while forgetting past data, according to newly read data based on a probability obtained by the probability calculation means, values of a mean parameter and a variance parameter of each normal distribution and a weighting parameter of each normal distribution, and

(Burge, especially: "Constructing Profiles")

degree of outlier calculation means for calculating and outputting a degree of outlier of said data by using a parameter of the normal mixture updated by said probability density estimation device and based on a probability distribution estimated from values of the parameters before and after the updating and the input data.

(Burge, especially: "The Fraud Engine")

26. In regards to Claim 4, Burge teaches the following limitations:

4. A probability density estimation device for use in a degree of outlier calculation device to, while sequentially reading a data sequence, estimate a probability distribution of generation of the data in question by using a finite number of normal kernel distributions, comprising:

parameter storage means for storing a value of a parameter indicative of a position of each kernel, and

(Burge, especially: "Constructing Profiles")

parameter rewriting means for reading a value of a parameter from the storage means and updating the stored parameter values while forgetting past data, according to newly read data to rewrite the contents of the parameter storage means.

(Burge, especially: "Constructing Profiles")

27. In regards to Claim 5, Burge teaches the following limitations:

5. A degree of outlier calculation device for sequentially calculating a degree of outlier of each data with a data sequence of real vector values as input, comprising:

a probability density estimation device for, while sequentially reading said data sequence, estimating a probability distribution of generation of the data in question by using a finite number of normal kernel distributions including
(Burge, especially: "Constructing Profiles")

(a) parameter storage means for storing a value of a parameter indicative of a position of each kernel, and
(Burge, especially: "Constructing Profiles")

(b) parameter rewriting means for reading a value of a parameter from the storage means and updating the stored parameter values while forgetting past data, according to newly read data to rewrite the contents of the parameter storage means, and
(Burge, especially: "Constructing Profiles")

degree of outlier calculation means for calculating and outputting a degree of outlier of said data by using said parameter updated by said probability density estimation device and based on a probability distribution estimated from values of the parameters before and after the updating and the input data.
(Burge, especially: "The Fraud Engine")

28. In regards to Claim 6, Burge teaches the following limitations:

6. For use in a degree of outlier calculation device for sequentially calculating a degree of outlier of each data with discrete value data as input, a histogram calculation device for calculating a parameter of a histogram with respect to said discrete value data sequentially input, comprising:

storage means for storing a parameter value of said histogram, and
(Burge, especially: "Constructing Profiles")

parameter updating means for reading said parameter value from the storage means and updating past parameter values while forgetting past data based on input data to rewrite the value of said storage means, thereby outputting some of parameter values of said storage means.
(Burge, especially: "Constructing Profiles")

29. In regards to Claim 7, Burge teaches the following limitations:

7. A degree of outlier calculation device for sequentially calculating a degree of outlier of each data with discrete value data as input, comprising:

a histogram calculation device for calculating a parameter of a histogram with respect to said discrete value data sequentially input including
(Burge, especially: "Constructing Profiles")

storage means for storing a parameter value of said histogram, and
(Burge, especially: "Constructing Profiles")

parameter updating means for reading said parameter value from the storage means and updating past parameter values while forgetting past data based on input data to rewrite the value of said storage means, thereby outputting some of parameter values of said storage means, and
(Burge, especially: "Constructing Profiles")

score calculation means for calculating, based on the output of the histogram calculation device and said input data, a score of the input data in question with respect to said histogram, thereby outputting the output of the score calculation means as a degree of outlier of said input data.
(Burge, especially: "The Fraud Engine")

30. In regards to Claim 8, Burge teaches the following limitations:

8. A degree of outlier calculation device for calculating a degree of outlier with respect to sequentially input data which is described both in a discrete value and a continuous value , comprising:

a histogram calculation device for estimating a histogram with respect to a discrete value data part,
(Burge, especially: "Constructing Profiles")

probability density estimation devices provided as many as the number of cells of said histogram for estimating a probability density with respect to a continuous value data part,
(Burge, especially: "Constructing Profiles")

cell determination means for determining to which cell of said histogram said discrete value data part belongs to send the continuous data part to the corresponding one of said probability density estimation

devices, and
(Burge, especially: “Constructing Profiles”)

score calculation means for calculating a score of said input data based on a probability distribution estimated from output values of said histogram calculation device and said probability density estimation device and said input data, thereby
(Burge, especially: “Constructing Profiles”)

outputting the output of the score calculation means as a degree of outlier of said input data,
(Burge, especially: “Constructing Profiles”)

said histogram calculation device including
(Burge, especially: “Constructing Profiles”)

storage means for storing a parameter value of said histogram, and
(Burge, especially: “Constructing Profiles”)

parameter updating means for reading said parameter value from the storage means and updating past parameter values while forgetting past data based on input data to rewrite the value of said storage means, thereby outputting some of parameter values of said storage means, and
(Burge, especially: “Constructing Profiles”)

said probability density estimation device including
(Burge, especially: “Constructing Profiles”)

parameter storage means for storing values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities and a weighting parameter of each normal distribution,
(Burge, especially: “Constructing Profiles”)

probability calculation means for calculating, based on a value of input data, and values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities, a probability of generation of the input data in question from each normal distribution, and
(Burge, especially: “Constructing Profiles”)

parameter rewriting means for updating and rewriting the stored parameter values while forgetting past data, according to newly read data based on a probability obtained by the probability calculation

means, values of a mean parameter and a variance parameter of each normal distribution and a weighting parameter of each normal distribution.
(Burge, especially: "Constructing Profiles")

31. In regards to Claim 9, Burge teaches the following limitations:

9. A degree of outlier calculation device for calculating a degree of outlier with respect to sequentially input data which is described both in a discrete value and a continuous value , comprising:
(Burge, especially: "Constructing Profiles")

a histogram calculation device for estimating a histogram with respect to said discrete value data part,
(Burge, especially: "Constructing Profiles")

probability density estimation devices provided as many as the number of cells of said histogram for estimating a probability density with respect to a continuous value data part,
(Burge, especially: "Constructing Profiles")

cell determination means for determining to which cell of the histogram said discrete value data part belongs to send the continuous data part to the corresponding one of said probability density estimation devices, and
(Burge, especially: "Constructing Profiles")

score calculation means for calculating a score of said input data based on a probability distribution estimated from output values of said histogram calculation device and said probability density estimation device and said input data, thereby
(Burge, especially: "Constructing Profiles")

outputting the output of the score calculation means as a degree of outlier of said input data,
(Burge, especially: "Constructing Profiles")

said histogram calculation device including
(Burge, especially: "Constructing Profiles")

storage means for storing a parameter value of said histogram, and
(Burge, especially: "Constructing Profiles")

parameter updating means for reading said parameter value from the storage means and updating past parameter values while forgetting past data based on

input data to rewrite the value of said storage means, thereby outputting some of parameter values of said storage means, and
(Burge, especially: "Constructing Profiles")

said probability density estimation device including
(Burge, especially: "Constructing Profiles")

parameter storage means for storing a value of a parameter indicative of a position of each kernel, and
(Burge, especially: "Constructing Profiles")

parameter rewriting means for reading a value of a parameter from the storage means and updating the stored parameter values while forgetting past data, according to newly read data to rewrite the contents of the parameter storage means.
(Burge, especially: "Constructing Profiles")

32. In regards to Claim 10, Burge teaches the following limitations:

10. For use in a degree of outlier calculation device for sequentially calculating a degree of outlier of each data with a data sequence of real vector values as input, a probability density estimation method of, while sequentially reading said data sequence, estimating a probability distribution of generation of the data in question by using a finite mixture of normal distributions, comprising the steps of:
(Burge, especially: "Constructing Profiles")

based on values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities read from parameter storage means for storing a value of input data, values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities, and a weighting parameter of each normal distribution, calculating a probability of generation of the input data in question from each normal distribution, and
(Burge, especially: "Constructing Profiles")

updating the stored parameter values while forgetting past data, according to newly read data based on a probability obtained by the probability calculation means, values of a mean parameter and a variance parameter of each normal distribution and a weighting parameter of each normal distribution to rewrite data of said parameter storage means.
(Burge, especially: "Constructing Profiles")

33. In regards to Claim 11, Burge teaches the following limitations:

11. A degree of outlier calculation method of sequentially calculating a degree of outlier of each data, with a data sequence of real vector values as input, wherein

probability density estimation for, while sequentially reading said data sequence, estimating a probability distribution of generation of the data in question by using a finite mixture of normal distributions, comprises the steps of:
(Burge, especially: "Constructing Profiles")

based on values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities read from parameter storage means for storing a value of input data, values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities, and a weighting parameter of each normal distribution, calculating a probability of generation of the input data in question from each normal distribution, and
(Burge, especially: "Constructing Profiles")

updating the stored parameter values while forgetting past data, according to newly read data based on a probability obtained by the probability calculation means, values of a mean parameter and a variance parameter of each normal distribution and a weighting parameter of each normal distribution to rewrite data of said parameter storage means, and which further comprises the step of:
(Burge, especially: "Constructing Profiles")

calculating and outputting a degree of outlier of said data by using a parameter of the finite mixture distribution updated by said probability density estimation and based on a probability distribution estimated from values of the parameters before and after the updating and the input data.
(Burge, especially: "The Fraud Engine")

34. In regards to Claim 12, Burge teaches the following limitations:

12. A probability density estimation method for use in calculation of a degree of outlier to, while sequentially reading a data sequence, estimate a probability distribution of generation of the data in question by using a finite number of normal kernel distributions, comprising the steps of:
(Burge, especially: "Constructing Profiles")

storing a value of a parameter indicative of a position of each kernel in parameter storage means, and reading

a value of a parameter from the storage means and updating the stored parameter values while forgetting past data, according to newly read data to rewrite the contents of the parameter storage means.
(Burge, especially: "Constructing Profiles")

35. In regards to Claim 13, Burge teaches the following limitations:

13. A degree of outlier calculation method of sequentially calculating a degree of outlier of each data, with a data sequence of real vector values as input, wherein

probability density estimation for, while sequentially reading said data sequence, estimating a probability distribution of generation of the data in question by using a finite number of normal kernel distributions comprises the steps of:
(Burge, especially: "Constructing Profiles")

storing a value of a parameter indicative of a position of each kernel in parameter storage means,
(Burge, especially: "Constructing Profiles")

reading a value of a parameter from the storage means and updating the stored parameter values while forgetting past data, according to newly read data to rewrite the contents of the parameter storage means, and which further comprises:
(Burge, especially: "Constructing Profiles")

degree of outlier calculation means for calculating and outputting a degree of outlier of said data by using said parameter updated by said probability density estimation and based on a probability distribution estimated from values of the parameters before and after the updating and the input data.
(Burge, especially: "The Fraud Engine")

36. In regards to Claim 14, Burge teaches the following limitations:

14. For use in calculation of a degree of outlier for sequentially calculating a degree of outlier of each data with discrete value data as input, a histogram calculation method of calculating a parameter of a histogram with respect to said discrete value data sequentially input, comprising the steps of:

reading said parameter value from storage means for storing a parameter value of said histogram and updating past parameter values while forgetting past data based on input data to rewrite the value of said

storage means, and
(Burge, especially: "Constructing Profiles")

outputting some of parameter values of said storage means.
(Burge, especially: "Constructing Profiles")

37. In regards to Claim 15, Burge teaches the following limitations:

15. A degree of outlier calculation device for sequentially calculating a degree of outlier of each data with discrete value data as input, comprising:

a histogram calculation device for calculating a parameter of a histogram with respect to said discrete value data sequentially input including
(Burge, especially: "Constructing Profiles")

storage means for storing a parameter value of said histogram, and
(Burge, especially: "Constructing Profiles")

parameter updating means for reading said parameter value from the storage means and updating past parameter values while forgetting past data based on input data to rewrite the value of said storage means, thereby outputting some of parameter values of said storage means, and
(Burge, especially: "Constructing Profiles")

score calculation means for calculating, based on the output of the histogram calculation device and said input data, a score of the input data in question with respect to said histogram, thereby outputting the score calculation result as a degree of outlier of said input data.
(Burge, especially: "The Fraud Engine")

38. In regards to Claim 16, Burge teaches the following limitations:

16. A degree of outlier calculation method of calculating a degree of outlier with respect to sequentially input data which is described both in a discrete value and a continuous value , wherein

histogram calculation which estimates a histogram with respect to a discrete value data part comprises the steps of
(Burge, especially: "Constructing Profiles")

reading said parameter value from storage means for storing a parameter value of said histogram and updating past parameter values while forgetting past data based on input data to rewrite the value of said storage means, and
(Burge, especially: "Constructing Profiles")

outputting some of parameter values of said storage means, and wherein
(Burge, especially: "Constructing Profiles")

in probability density estimation devices provided as many as the number of cells of said histogram for estimating a probability density with respect to a continuous value data part, said method comprises the steps of:
(Burge, especially: "Constructing Profiles")

based on values of a mean parameter and a variance parameter of each of a finite number of normal distribution densities read from parameter storage means for storing a value of input data, values of a mean parameter and variance parameter of each of a finite number of normal distribution densities and a weighting parameter of each normal distribution, calculating a probability of generation of the input data in question from each normal distribution, and
(Burge, especially: "Constructing Profiles")

based on a probability obtained by the probability calculation means, values of a mean parameter and a variance parameter of each normal distribution and a weighting parameter of each normal distribution, updating the stored parameter values while forgetting past data, according to newly read data to rewrite the data of said parameter storage means, and wherein said method further comprises the steps of:
(Burge, especially: "Constructing Profiles")

determining to which cell of said histogram said discrete value data part belongs to send the continuous data part to the corresponding one of said probability density estimation devices,
(Burge, especially: "Constructing Profiles")

calculating a score of said input data based on a probability distribution estimated from output values of said histogram calculation device and said probability density estimation device and said input data, and
(Burge, especially: "The Fraud Engine")

outputting the score calculation result as a degree of outlier of said input data.
(Burge, especially: "The Fraud Engine")

Conclusion

39. The Yamanishi et al. reference (See p.320, col.2, para. 3) teaches the following about the Burge reference:

> Note that there exists only a few works (e.g. Burge) focusing on the on-line unsupervised learning based approach [to outlier detection in data mining].

and also the following (See p.321, col.1, para.5) about the Burge reference:

> The design of SS [SmartSifter] was inspired by the work by Burge and Shawe-Taylor. Our work differs from [Burge] in the following regards:
>
> 1) SS [SmartSifter] treats both categorical and continuous variables, while [Burge] deals only with continuous ones.
>
> 2) While Burge uses two models in the algorithm: the long term model and the short term one, SS [SmartSifter] unifies them into one model with the aim of a clearer statistical meaning and a lower computational cost.
>
> 3) SS [SmartSifter] uses either a parametric representation for a probabilistic model or a non-parametric one, while only a non-parametric one is used in [Burge]. In Sec.3.1, we compare our parametric method with the non-parametric one to show that the former outperforms the latter both in accuracy and computation costs.

40. Applicants are reminded that these features differentiate the current application from the prior art, and should be in the claims.

Correspondence Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ayal I. Sharon whose telephone number is

(703) 306-0297. The examiner can normally be reached on Monday through Thursday, and the first Friday of a biweek, 8:30 am – 5:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Kevin Teska can be reached on (703) 305-9704. Any response to this office action should be mailed to:

Director of Patents and Trademarks
Washington, DC 20231

Hand-delivered responses should be brought to the following office:

4th floor receptionist's office
Crystal Park 2
2121 Crystal Drive
Arlington, VA 22202

The fax phone number is: (703) 872-9306

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist, whose telephone number is: (703) 305-3900.

Ayal I. Sharon

Art Unit 2123

May 14, 2004

KEVIN J. TESKA
SUPERVISORY
PATENT EXAMINER